DEPARTMENT OF THE NAVY
NAVY PERSONNEL RESEARCH AND DEVELOPMENT CENTER
SAN DIEGO, CALIFORNIA 92152-6800

3900
Ser 411/815
9 OCT 1987

From: Commanding Officer, Navy Personnel Research and Development Center

Subj: **METHODS OF ELICITING INFORMATION FROM EXPERTS**

Encl: (1) NPRDC TN 88-2

1. The work presented in enclosure (1) was performed to consider the methods available for extracting information from experts. The report is intended primarily for those within the scientific community who are engaged in conducting research and development in the area of artificial intelligence and expert systems.

3. The point of contact at the Navy Personnel Research and Development Center is Dr. David Meister, (619) 225-2081 or AUTOVON 933-2081.

RICHARD C. SORENSON
By direction

Distribution:
Office of Naval Research (Code OCNR-1142)
Office of Naval Technology (Code 222)
Defense Technical Information Center (2)

NPRDC TN 88-2 October 1987

METHODS OF ELICITING INFORMATION FROM EXPERTS

David Meister

Navy Personnel Research and Development Center
San Diego, California 92152-6800

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Navy Personnel Research and Development Center 6b. OFFICE SYMBOL (If applicable) Code 41	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) San Diego, CA 92152-6800	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Methods of Eliciting Information from Experts

12. PERSONAL AUTHOR(S)
Meister, D.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Note	FROM 1987 TO 1987	1987 October	26

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Expert systems, human expertise, knowledge acquisition, mental processes, system design
05	01		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The biggest bottleneck in the development of expert systems is the problem of eliciting from experts the mechanisms responsible for their expertise. This report examines what is known about experts and suggests a number of ways of eliciting information from them.

The literature suggests that the mechanisms of expertise represent deep-seated ways of conceptualizing and perceiving stimuli, and that these mechanisms must be differentiated from relatively superficial procedura rules, which make up most of what "expert information" consists of today.

The goal of reproducing the expert's mental processes in a computer system appears unrealistic at present. The only way of determining that one has, in fact, tapped expertise is to build the expert system and evaluate its effectiveness. If it matches or surpasses human proficiency, one has incorporated human expertise into the system.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL D. Meister	22b. TELEPHONE (Include Area Code) (619) 225-2081	22c. OFFICE SYMBOL Code 41

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

SUMMARY

The biggest bottleneck in the development of expert systems is the problem of eliciting from experts the mechanisms responsible for their expertise. This report examines what is known about experts and suggests a number of ways of eliciting information from them.

The literature suggests that the mechanisms of expertise represent deep-seated ways of conceptualizing and perceiving stimuli, and that these mechanisms must be differentiated from relatively superficial procedural rules, which make up most of what "expert information" consists of today.

The goal of reproducing the expert's mental processes in a computer system appears unrealistic at present. The only way of determining that one has, in fact, tapped expertise is to build the expert system and evaluate its effectiveness. If it matches or surpasses human proficiency, one has incorporated human expertise into the system.

CONTENTS

INTRODUCTION

Purpose

The purpose of this report is to examine and compare ways in which one can elicit from an expert a knowledge of the mechanisms on which the expertise is based.

Knowledge acquisition for expert systems is a difficult and time-consuming process, the biggest bottleneck in the production of these systems. Unfortunately, very little is known about how to extract expertise from experts, and almost nothing is available as a reliable technique. On the one hand, one has experts who are unfamiliar with expert systems and relatively inarticulate about the expertise they possess, and how they use it to solve problems; and, on the other hand, one has a knowledge engineer who, in most cases, is ignorant about the knowledge domain at issue and is almost never skilled in behavioral techniques.

Expert information is desired in connection with the development of so-called expert computer systems. These systems attempt to incorporate the mechanisms used by experts in performing those functions at which they are considered "expert." Some specialists in artificial intelligence even talk about reproducing in the software the actual mental processes used by experts. Whether this is a reasonable goal or not (and a later discussion will suggest that it is not), it provides the rationale for attempting to elicit as much information as one can from an expert. Unfortunately, experts vary in their ability to communicate the "secrets" of their expertise. The reason for this may be that the mechanisms responsible for their expertise are essentially unconscious.

A distinction must be made between the information or knowledge available to an expert and the mechanisms responsible for that expertise. The two are not the same. As an illustration, it is possible to be very knowledgeable about football-- many fans are--but not to be a star football player. Although both the football superstar and the football fan may possess the same information about football, only the former possesses true expertise. Is the information supplied by the superstar any more accurate than that provided by the fan? If the true expertise mechanisms cannot readily be elicited, does it make any difference who is asked for procedural information, presuming that the fan really knows the game? Knowledge engineers must ask themselves whether they are attempting to elicit procedural knowledge or mechanisms. It is possible to confuse the two. Certainly what is most immediately available to the knowledge engineer is knowledge. This report makes a judgment on this point, which, however, requires examination.

A number of questions are inherent in the process of eliciting information from experts and incorporating it into software:

1. Are the mechanisms responsible for expertise accessible to experts? To other than experts? If they are not accessible, then obviously they cannot be retrieved or retrieved only with the greatest difficulty.

2. Assuming that these mechanisms can be retrieved, are they translatable into algorithmic form? Are there computer languages suitable to describe these mechanisms? In order to incorporate expertise into a computer system, it is necessary to translate these mechanisms into proceduralized form. This may not be easy to do if the mechanisms are not procedural in nature.

3. Is there any way of knowing when one has secured sufficient data from experts to make expert systems effective? If an objective criterion of information sufficiency is unavailable, one can not be sure that one has secured sufficient knowledge until after the expert system is built and tested.

A negative answer to any one of these questions casts serious doubt on the feasibility of incorporating human expertise into expert systems.

Who is an Expert?

Before focusing on the problems raised earlier, one preliminary question must be answered. Who is an expert, how does one identify him or her? In some cases, those in which there is no external criterion, the judgment must be subjective and therefore must be consensual (i.e., a matter of agreement among judges). In other cases the judgment can be related to performance (i.e., winners of Olympic medals or Nobel Laureates are automatically expert in their domains).

The question of identifying the expert is one that cannot be answered glibly, because there are different types of experts, and some types may be easier to elicit information from than others. Then, too, if one identifies someone as an expert who is not actually an expert, the information elicited may be erroneous or inadequate for developing an expert system.

Experts may be classified in terms of the predominant behavioral function involved in their expertise. Some expertise is predominantly psychomotor, as in the case of sports figures or dancers. Other expertise may be perceptually oriented, as in the case of an artist. A combination of modalities may be involved, for example, the musician who has sensory expertise (e.g., perfect pitch) as well as psychomotor expertise (a violinist's bowing). Other expertise may be primarily cognitive, as in the case of a medical diagnostician.

It may be easier to elicit the underlying mechanisms when the expertise is cognitive than when it is psychomotor, when it involves a single function rather than several combined. There have been attempts in the past, for example, to determine what the basis of fighter ace expertise is, without success, probably because the researchers could not elicit the relevant information. There are degrees of expertise. One might be a violin virtuoso (e.g., Itzhak Perlman) or a concertmeister in an orchestra, or simply one of its violinists. Differences in amount of expertise may supply different expertise mechanisms and may affect the ease with which one can elicit information.

In general, expert system developers have attempted to make use of cognitively oriented experts because the systems for which developers were responsible were designed to perform cognitive functions (e.g., medical diagnosis (INTERNIST) or geological analysis (PROSPECTOR)). So far government and industry have not been interested in developing an expert system that would lead to becoming a world class violinist.

Underlying Assumptions

In attempting to elicit information from experts, one makes a number of assumptions: (1) that expertise mechanisms are communicable to others, that is, that these mechanisms are translatable into what we ordinarily call information; (2) that an expert is aware of or can be stimulated to become aware of these mechanisms; (3) and that these expertise mechanisms (which are ordinarily covert) are related to symbolic concepts that can be

expressed verbally and/or graphically. The relationship between mechanism and concept may or may not be close.

Since the mechanisms responsible for the expertise reside in the expert and cannot be observed or derived except through the medium of the expert, the expert must be involved in the examination of these mechanisms. The effects of these mechanisms can be observed in the expert's performance outputs, but the outputs are not the same as the mechanisms. In considering these outputs all one can do is to infer certain qualities of the mechanism based on the qualities of the output. Also, experts must be able to observe these mechanisms or they cannot reveal them. When an expert says that what is revealed as the mechanisms is completely faithful to those mechanisms, the observer has a choice of believing the expert or not. So much subjectivity is involved that one can be certain of nothing. Manifestly there is no external criterion of validity.

The assumptions present an appalling case for the procedure of eliciting expertise, but they can be ignored. The ultimate proof of the pudding, as it were, is the expert system itself. If it does as well as or better than the expert's expertise, then one can assume that the elicitation process has been validated. However, this is only an assumption like the previous ones. It is conceivable that the completed expert system was made proficient by mechanisms (e.g., the developer's rules of thumb, logic, inspired engineering, guesswork) other than those derived from the expert. Even in performance terms one cannot be quite certain the elicited expertise was responsible for system success. All one knows is that something--as yet unknown--produced an effective expert system. If the new system does not do as well as the developer hopes, it cannot be said that the system developer misinterpreted the expert, only that something was responsible for system failure.

The preceding should not be interpreted to mean that developers should ignore human expertise in development of the expert system, or should not try their damndest to get as much as they can out of experts. The correct interpretation is that if an expert is the basis for the system, the developer is building the system on very tenuous foundations. The solution to this problem is to consider human expertise as only one of a number of inputs to system development. Complete reliance on expertise poses the danger that the expertise may be inadequate.

How can one tell when the expert's product, a description of how to do something, actually represents that activity? The knowledge engineer looks at the verbal product and compares this with a vague concept of what the output should be.

If the one eliciting the expertise from an expert is likewise an expert, the chances are much improved of getting a useful product, because the standard by which one can evaluate the product exists in the second expert. If one great pianist interrogates another pianist about how the latter achieves pianistic effects, the interrogator, being an expert too, is unlikely to be "snowed" by what the subject has to say.

The chances of getting a useful product are also enhanced if several experts are used as subjects for the interrogation. The expertise of any single one may rest on purely idiosyncratic grounds that are not translatable into algorithms.

Even if one assumes that the product genuinely reflects "true" expertise, it must be translated into software and here that translation may be poor, negating the "truth" of the original product. If, after eliciting the product as a verbal protocol, one takes this protocol to an expert and asks, "Is this how you do it?", the expert may not be able to

recognize his or her expertise in the protocol and, thus, deny it. Or the expert may agree, quite incorrectly. The expert is his or her own supreme arbiter of a product's adequacy in representing that expertise. It would be delightful to try to determine the characteristics of product adequacy, but no one has investigated this, though there may, in fact, be some evidence.

An indicator of an adequate product might (the following is hypothesis only) include the amount of detail and/or amount of material provided. The more detail, the more likely the protocol is adequate. The inclusion of quantitative data in the product (if such data are relevant to the area of expertise) might be another potential index. One can say nothing more about these indices, because to our knowledge no one has ever studied the question empirically, and empirical investigation is what is needed.

Another potential criterion of product adequacy might be the ease with which the product is transformed into algorithms and software. The product might appear to be more adequate to the knowledge engineer if it is logical and internally consistent.

All such criteria have inherent deficiencies. An expert protocol may be logical or be easily translatable into software only because essential elements of the expert's mechanisms have not been elicited in the protocol.

One criterion that is highly attractive is the securing of approval by a second expert. A musician should be able not only to ask the appropriate questions but also to evaluate the adequacy of the replies received.

Implicit in all this are certain assumptions central to the elicitation process:

1. One must guide experts (even badger them) to provide relevant and comprehensive data.

2. Experts' initial products will be unsatisfactory and will have to be refined in successive trials.

3. Experts vary in degree of expertise and ability to express themselves.

It is not expected that experts will be able fully to recognize the mechanisms of their expertise or be able to organize verbalizations that are fully intelligible or be able to communicate fully. It will be necessary to stimulate experts by asking questions in much the same way that a lawyer stimulates a witness. But does this not require the interrogator to have some expertise in the knowledge domain at issue?

The individual who does this and who turns (or at least helps to turn) the product into software algorithms is known as the knowledge engineer. There has been no discussion as to what special capabilities this individual should have, but since the elicitation of information from an expert is probably an interactive situation, some attention should also be paid to the qualities the engineer should have.

Manifestly the task of eliciting information from experts is a task fraught with difficulty and danger. Experts may provide inadequate detail or may (although this is somewhat unlikely) be incorrect in some aspect of what they say; the products may be contaminated by idiosyncrasies that are irrelevant to the expertise; some of the material may be irrelevant.

Are there formal procedures for eliciting information from experts? Probably not. It seems reasonable, however, to believe that if the expertise is manifested in performance of a task, the questions asked of experts should also be oriented around task performance. The following questions are generic in the sense that they can be asked about any expertise domain.

1. What are the elements to be considered in performing the expert task?

2. What are the interrelationships among the elements, i.e., dependencies?

3. What are the stimulus cues the expert is responsive to?

4. What specific procedures must be performed? What are the initiating conditions for these?

5. What is the sequencing of these procedures?

6. What outputs of performing the expertise task should one expect?

7. What constrains expert task performance?

8. What conditions must be taken into account in performing the expert task?

9. What task-consequence relationships exist, that is, if I do this, what will result?

Characteristics of the Expert

What the Literature Tells Us

The literature on experts does not, except incidentally, deal with the problem of eliciting information from experts. The focus of research attention is on the question of how experts differ from nonexperts (i.e., What mechanisms do they employ that make them experts or are associated with the expertise and that are not to be found in the non-expert?). The answers to this question may suggest solutions to the elicitational problem, but only indirectly. However, since this is the only expert literature extant, it is necessary to consider it, although not in as much detail as one would if the central question of that literature were the subject of this report.

The paradigm in investigations of expert behavior is to identify an expert by whatever criteria are acceptable to the researcher; to contrast the expert with another subject who is definable as a nonexpert (because he differs from the expert); and then to present both groups of subjects with the same problem. As each subject solves the problem, he or she verbalizes thought processes. The verbal protocol, together with other evidence such as performance observations, time to solve, percent success in solving the problem, etc., is analyzed to reveal the mechanisms that differentiate the two sets of subjects.

For example, Johnson, Hassebrock, Duran and Moller (1982) had expert and novice clinicians judge the likelihood of disease alternatives and provide thinking-aloud protocols as they evaluated simulated cases of congenital heart disease. Combinations of cues in the patient data were manipulated to create alternative versions of a single case so that the use of critical cues could be identified. This is an illustration of a policy-capturing methodology that can be used to identify the cues experts respond to. Unfortunately one

of the puzzling findings from studies of expert-novice differences in problem-solving behavior is that differences between experts with regard to the means used to solve a given problem are often as great as the differences between experts and novices (Chase & Simon 1973; Elstein, Shulman, & Sprafka, 1978).

The literature (see the following extract from Adelson, 1984) suggests that the mechanisms that differentiate experts from nonexperts are much more fundamental than what might be presumed from verbal expression of procedures used in applying one's expertise.

> For example, in skilled problem solving (Bhaskar & Simon, 1977), the difference between experts and nonexperts is both qualitative and quantitative. Not only do experts perform better than novices on quantitative measures of skill but experimental manipulations also uncover qualitative differences in the representations and strategies used by experts. Repeatedly, we find that the working representations of experts are abstract conceptualizations of the original problem statement, whereas those of novices are less abstract and focus more on surface features of the problem. For example, in a recent experiment, Adelson (1981) found that expert programmers used abstract, conceptually based representations when attempting to recall programming material, whereas novices used syntactically based representations. Using a multitrial free-recall procedure, Adelson asked novice and expert programmers to recall a set of 16 lines of programming code that had been presented in random order. Although the subjects had not been told that the 16 lines could be organized either conceptually into three programs or syntactically into five categories according to the control words that they contained, analyses of the order of recall for each group showed that the experts had clustered the lines into complete programs, and the novices had clustered the lines according to syntactic categories.
>
> The classic result on the abstract nature of the representations of experts was obtained by Chase and Simon (1973). They replicated de Groot's (1965) findings in which Master chess players reconstructed with greater than 90% accuracy midgame boards that they had seen for only 5 s. They then went on to isolate and to characterize the chess Masters' recall clusters and found that clusters frequently consisted of chess pieces that formed attack or defense configurations. This observation suggests that individual chess pieces are seen as integral parts of larger, meaningful units. Looking at the recall clusters of Master Go players, Reitman (1976) also found abstract representations that were based on the attack and defense relationships in the game board. McKeithen, Reitman, Rueter, and Hirtle (1981) found that intermediate programmers cluster the words of a programming language by concept, whereas beginners cluster the same words alphabetically. Schneiderman's (1977) finding that the recall distortions of skilled computer programmers preserve the concepts but not the specific form of previously seen material also suggests an abstract representation.
>
> Chi, Glaser, and Reese (1981) have drawn inferences about the schemata of novice and expert physicists from the results of

conceptual sorting tasks and verbal protocols. They suggested that the schema of the novice represents the surface features of the problem, whereas the schema of the expert represents the abstract physical principles involved plus conditions that specify when to apply the principles.

Lewis (1981) examined the solutions of experts and novices in algebra problems. He found that experts often restructure the terms in the original problem, but novices never do. The kind of restructuring that Lewis found suggests an abstraction of the elements of the problem.

Taken together, the above findings seem to support the suggestion that experts form abstract, conceptual representations of problems but novices form representations that tend to retain the surface elements of the problem.

The verbal protocol recording is the preferred method for examining problem-solving processes; this is demonstrated by the number of studies that have used it. Following the pioneering work of Newell and Simon (1972) in cryptarithmetic, it has been used in a variety of domains: physics (Simon & Simon, 1978; Larkin, McDermott, Simon & Simon, 1980; Larkin, 1981; Chi, Feltovich & Glaser, 1981), mathematics (Anderson, Greeno, Kline & Neves, 1981; Lewis, 1981), financial analysis (Bouwman, 1978, 1983; Biggs, 1978a, b), software design (Malhotra, Thomas, Carroll & Miller, 1980; Jeffries, Turner, Polson, & Atwood, 1981), and systems analysis (Vitalari & Dickson, 1983).

However, sometimes the techniques used to explore expert structure are quite abstract and even mathematical. Schvaneveldt et al. (1985) used multidimensional scaling and link-weighted networks and suggested the possible uses of these techniques to elicit expert knowledge in a form appropriate for coding into assertions and rules. In their study, subjects (fighter pilots) rated the similarity of relationship of 435 pairs of terms (i.e., 30 basic concepts taken two at a time). The obtained similarity measures were transformed into measures of psychological distance by subtracting the ratings from the maximum possible rating. The data for each subject (a 30 x 30 symmetrical matrix) were exposed to multidimensional scaling procedures.

One might suppose that knowledge about the dimensions of expertise might assist in suggesting ways of eliciting information from experts. Glaser (1985) has summarized the results of studies by himself and others concerning the nature of expertise.

Studies of problem solving in adult experts and novices have shown fairly consistent findings in a variety of domains--chess play, physics problem solving, the performance of architects and electronic technicians, and skilled radiologists interpreting x-rays. This work has shown that relations between the structure of a knowledge base and problem-solving processes are mediated through the subject's representation of the problem. This problem representation is constructed by the solver on the basis of domain-related knowledge and the organization of this knowledge. The nature of this organization determines the quality, completeness, and coherence of the internal representation, which, in turn, determines the efficiency of further thinking.

Expert/novice research suggests that novices' representations are organized around the literal objects and events given explicitly in a problem statement. The expert's knowledge, on the other hand, is organized around inferences about principles and abstractions that subsume these factors. These principles are not apparent in the statement or the surface presentation of the problem. For example, in studies with mechanics problems, novices classify problems on a surface level, in terms of the physical properties of a situation--a spring problem or an inclined plane problem. Experts categorize problems at a higher level, in terms of applicable physics principles--a Newton's second law problem, a conservation of energy problem.

In addition, experts know about the application of their knowledge, which is tightly bound to conditions and procedures for its use. A novice may have sufficient knowledge about a problem situation, but lack knowledge about conditions of applicability.

Specific Findings

The following generalizations from Glaser (1985) seem reasonable:

1. There seems to be a continuous development of competence as experience in a field accumulates. Eventual declines in competence may be the result of factors in an expert's experiences. Competence may be limited by the environment in which it is exercised. People may attain a level of competence only insofar as it is necessary to carry out the activities or to solve problems at the given level of complexity presented.

2. Expertise seems to be very specific. Expertise in one domain is no guarantee of expertise in other areas. It may be, however, that certain task domains are more generalizable than others, so that those who are experts in applied mathematics or aesthetic design have forms of transferable expertise.

3. Experts develop the ability to perceive large meaningful patterns. These patterns are seen in the course of their everyday activities. This pattern recognition occurs so rapidly that they take on the character of "intuitions." In contrast, the patterns that novices recognize are smaller, less articulated, more literal and surface-oriented, and much less related to inferences and abstracted principles.

4. Experts' knowledge is highly procedural. Concepts are bound to procedures for their application and to conditions under which these procedures are useful. Experts' functional knowledge is related strongly to their knowledge of the goal structure of a problem. Experts and novices may be equally competent at recalling small specific items of domain-related information, but high-knowledge individuals are much better at relating these events in cause-and-effect sequences that relate to the goal and subgoals of problem solution.

5. These components of expertise enable fast-access pattern recognition and representational capability that facilitate problem perception in a way that greatly reduces the role of memory search and general processing. Novices, on the other hand, display a good deal of search and processing of a general nature. Their perceptions are highly literal and qualitatively different from experts' representations. Experts and novices work with similar capacity for processing; experts' outstanding performance derives from how their knowledge is structured for processing.

In the course of acquiring expertise, novices experience plateaus of development that appear to indicate shifts in understanding and stabilizations of automaticity (unconscious processing).

Expert representations of problems and situations are qualitatively different from novice representations. In the course of a novice developing expertise, problem representation changes from surface representations to inferred problem descriptions, to principled (and proceduralized) categorizations.

In some domains, experts are "opportunistic planners"; new problem features result in changed problem representation. They show fast access to multiple interpretations; novices are less flexible (e.g., x-ray and medical diagnosis, Lesgold, Feltovich, Glaser, & Wang, 1981).

Experts can be disarmed by random (or meaningless) patterns and lose their great perceptual ability. (For example, with a scrambled chessboard, experts and novices do equally poorly.)

Experts are schema-specialized and these schemata drive their performance. (Experts impose a structure on a noisy x-ray; novices are misled by this noise.)

They are goal-driven: Given a complex goal, they will represent the problem accordingly; given simple goals, they will think only as deeply as necessary.

Experts display specific domain intelligence, not necessarily general intelligence.

Novices make extensive use of general heuristic problem-solving processes (of the Newell and Simon variety, e.g., generation and testing, means-end analysis, subgoal decomposition); experts use them primarily in unfamiliar situations.

Experts may be slower than novices in initial problem encoding but are overall faster problem solvers (e.g., analogical reasoning test items, Sternberg, 1977a).

The development of expertise is subject to task demands and the "social structure" of the job situation; the cognitive models that experts acquire are constrained by task requirements (e.g., Scribner, 1984).

Expertise in some knowledge domains may be more generalizable (broadly applicable) than that in other domains.

Experts develop automaticity, particularly of "basic operations," so that memory is available for conscious processing.

In solving ill-structured problems, experts employ relatively general methods of problem decomposition, subgoal conversion, and single factor analysis; their initial thinking is less driven by principles and procedural aspects related to their specific knowledge domain. Experts work from their memory of an issue's history to represent problems and devise arguments for alternative solutions (e.g., see analysis by political scientists (Voss, Green, Post, & Penner, 1983)).

Experts become skilled in the use of self-regulatory processes such as solution monitoring, allocation of attention, and sensitivity to informational feedback (Brown, 1978).

Expert knowledge is not inert; it is highly proceduralized and conditionalized (Anderson, 1983).

Super experts may develop generalizable abilities through the use of mapping and analogy (Gentner & Gentner, 1983).

General thinking and problem-solving skills may develop in the course of shifting between many domains, so that the cognitive processes involved become decontextualized (Glaser, 1984).

What the preceding statements suggest is that (as one might suspect) mechanisms of expertise are deeply buried in an expert's consciousness and not necessarily tied to any factual (easily retrievable) information. It is significant that the preceding conclusions did not refer to conditions of eliciting information from experts. The review of the literature suggests that methods of eliciting data from experts are highly limited in the number of options possible. One can interview experts, observe and evaluate their performance, and examine the products of their work--and that is all.

Computerized Methods

Most of the methods of eliciting information from experts are manual, but efforts have been made to computerize the process. Prototype knowledge acquisition systems that interview experts have been built to support several expert system problem areas (Boose, 1986).

> MDIS (Antonelli, 1983) and MORE (Kahn, Nowlan & McDermott, 1985) are capable of eliciting more sophisticated responses by including domain knowledge during interviewing. MDIS leads in a structured breakdown of a mechanism and elicits causal relationships between modules. The relationships may be descriptive or functional. MDIS analyzes these descriptions and classifies sets of modules into higher level objects; these objects then become keys to performing more efficient diagnosis. Several types of diagnostic strategies (such as failure rates and causal relationships among modules) are taken into account, and the system generates sets of production rules for diagnosis. MORE builds diagnostic models based on links between hypotheses, symptoms, and tests. The strategy of having the expert make distinctions between these objects where necessary drives the knowledge expansion process and fills in gaps in the model.
>
> These systems work on the structured selection portion of the problem spectrum. They have their origins in a system called TEIRESIAS (Davis & Lenat, 1982). TEIRESIAS helps experts incrementally refine the knowledge base. It helps debug rules based on specific cases and by building models of knowledge bases in progress. TEIRESIAS can tell whether a new rule fits the current model for that particular type of rule and can even suggest new rules.

SALT (Marcus, McDermott & Wang, 1985) is a system that interviews experts to help in configuring systems (it was first used to configure elevators). In some sense it is the first knowledge acquisition interviewing system that bridges the gap between analysis and synthesis problems. The approach used by SALT is successful to the extent that the expert can define the search network. SALT elicits knowledge about values of specific configuration parts, relationships between parts, and recognition and remedy of constraint violations. SALT's developers are also planning to apply it to scheduling problems.

Acquiring knowledge for general planning problems is much more difficult than acquiring knowledge for configuration tasks since the portion of the search network that can be explicitly enumerated is usually small.

PLANET, another interviewing program based on methods from personal construct psychology (Shaw, 1982), combines ideas from system theory, psychology, and application methodologies.

Boose (1986) has developed something called the expertise transfer system (ETS), which is derived from George Kelly's (1955) personal construct theory. ETS uses the technique to help experts explore the way they solve problems. The methodology has been applied to the task of eliciting information from travel agents. The information is gathered by asking an expert to distinguish between objects and classes of objects.

All of these methods are in experimental prototype form only and no final word can be given about their success. It is not known, for example, how well such systems "stack up" against manually elicited expert outputs. It should be noted that special software would have to be developed for each knowledge domain, an effort that may be somewhat daunting. The automated knowledge acquisition system is a general-purpose system only to the extent that general principles are the foundation of the software used to explore a particular knowledge domain. If the knowledge domain changes, the general principles must be retailored to the new domain. Developing such a knowledge acquisition expert system may be even more complex than gathering the same information manually. Indeed, one has to gather that information manually before one can develop the knowledge acquisition system, so that there appears to be little net value in using the system. Only if the general-purpose knowledge acquisition system could be used in exploring a new knowledge domain with only minor modification to its software would the value of an expert knowledge acquisition system be demonstrated.

MANUAL METHODS OF INFORMATION ACQUISITION

In the acquisition of information from experts, several of the following techniques may be combined on the principle that each technique may elicit information of a somewhat different nature.

Interviewing

The most common method is to interview experts about the principles they use to solve problems and to make diagnoses or decisions. The interview may be a general one, to elicit introductory material or to follow up on a problem solution. A representative

problem may be presented either verbally or in written/graphic form as a context for the interview (e.g., transparencies of cell structure for a biological diagnosis). The expert may be asked the following questions, which are general in nature.

1. Are all the problems you solve much the same, at least are they in the same general form? If not, how do they differ? What effect do the differences have on the way in which you approach the problem?

2. How do you start the process in which your expertise is manifested? How would you characteristically describe that process, if you had to choose one or two words? Are there recognizable stages in that process? If so, what differentiates one stage from another?

3. How do you conceptualize the problem you are faced with? What are the elements of the process or problem? What cues would you pay most attention to? What information do you need? Does your information need change during the process? What principles or rules do you use to direct your work? What factors (e.g., cues, information) do you trade off? Given cues X, Y and Z, what are the interrelationships among them? Are any cues or any pieces of information particularly important (have greater influence on your decisions)? (It might be useful for the knowledge engineer to plot the cues and information sources graphically and attempt to show interrelationships by drawing lines in the manner of a link analysis. This could be shown to the expert later and the expert asked to rate the strength of the interrelationships plotted.)

4. What factors in the problem are dependent on other factors? How strong are the dependencies among factors? Is there a finite number of problem solutions, and what are they? How do you decide which one of the solutions is best? What kind of information verifies or refines that solution?

The questions above focus on the cues the expert responds to and the interrelationship of elements in the process. That is because it is assumed that regardless of the specific nature of the process, in a general sense, it is a perceptual/cognitive problem and the expert's responses to these cues go to the heart of his or her expertise.

These are generic questions. Manifestly they would have to be tailored more specifically to a particular knowledge domain. Moreover, if they were asked following a demonstration by an expert of his or her expertise, the questions would more directly concern the actions taken by that expert.

Problem Solution

Another common method is to present experts with one or more problems by which to demonstrate their expertise. A single problem is not adequate, since any single problem may inadvertently call forth only special mechanisms. In the ideal situation, problems would differ in terms of a set of predetermined dimensions so that one could test the effect of the dimensions on the expert's process. This assumes that the knowledge engineer has enough domain knowledge (but still would almost certainly require the aid of an expert consultant) and the problems are sufficiently clear-cut and discrete that the dimensions underlying them can be ascertained in advance. The keynote of these techniques is that the correct answer to the problem posed is already known (or at least highly suspected) and the ways in which the problems differ are also known.

One suspects that in most cases the above conditions will not exist and at best the problems presented will be considered "representative" (also determined by consultation with one or more experts), and the knowledge engineer will have only partial control over the situation. The expert solves the problems and the performance is observed and recorded, but the expert does not verbalize during the solution. Following the solution of each problem the expert is interviewed concerning the methods employed. The preceding interviews are utilized, but tailored to the individual problem. Emphasis is placed on why the expert did what he or she did. The investigator reviews the process of problem solution in detail, at major points asking the expert to tell what was done, why it was done, and what information was used or collected. The knowledge engineer's notes and observations are used to stimulate the expert during the post-solution interview.

The reason for having experts actually solve problems or whatever else they do in utilizing their expertise is to see the expertise mechanisms functioning dynamically. The method is deficient in that experts may not remember precisely what they did (although, unless the solution process is excessively slow, this should not be a serious factor). One might videotape the problem solution process, so that it can be played back to the expert while questioning proceeds. This, of course, assumes that the process involves physical manipulations; one cannot videotape covert processes.

Verbalization

Closely related to the preceding methods is one requiring experts to verbalize what they are thinking and/or doing during the solution of the problems presented to them. Experts must be trained to verbalize or at least told the categories of knowledge in which one is interested. One cannot be cavalier about this aspect; most experimental subjects do not verbalize adequately without training or coaching. The categories explored will be roughly the same as those used in interviewing, with the addition that one wishes to know <u>why</u> experts are doing what they are doing during the solution process. In the most effective form of this situation, experts are presented with a series of problems to solve. They are told what topics they should verbalize about and then respond as requested. They are observed during problem solution (perhaps using a videotape as well) and, following solution, intensively interrogated by the knowledge engineer. For obvious reasons it is best if the verbalization is tape-recorded; it can then be played back to experts as part of the interview.

It should be noted that the requirement to verbalize while solving problems may inhibit the expert's solution process (to what degree is not known); some experts may be more susceptible to this blocking than others.

Verbalization without interviewing poses the risk that experts may be deficient in communicating and verbalizations may be shallow and nonproductive. Intensive post-solution interviewing will compensate for this, if, in fact, experts are deficient in verbalization and/or self-awareness. Some people are, in fact, experts but are unaware of the mechanisms that make them experts. Investigation of an expert's fluency and self-awareness is recommended before an individual is accepted as a subject expert, even though the individual is <u>technically</u> qualified.

The ideal arrangement would be for each expert to serve as an observer/interrogator of other experts, because the former is most likely to be aware of the quality of an expert's responses and can stimulate these during the follow-on interview. It is assumed that more than one expert will be used as a subject; individual differences among experts are well known.

A variant of the free flow verbalization method is to allow experts to verbalize freely and then to stop them at convenient points in the problem solution to ask penetrating questions. Such a procedure runs the risk of inhibiting the verbalization flow and should be carefully considered before being implemented. The method has the advantage of securing more detailed data from the expert than would otherwise be gained.

Checklist

One technique that might prove useful is to give experts lists of possible dimensions, factors, or cues that could influence judgment and have them check one or more off at critical points (to be specified) during the problem solution. Such a checklist serves to structure responses. However, such a checklist depends on prior work for its development; this may not be simple because it requires some expertise on the part of the checklist developer. However, one or more experts (working independently or as a group) could be used to develop such a checklist; indeed, the checklist development effort might be an excellent means of securing information from a group of experts (during its development, that is). An alternative way of securing expert information is to develop a questionnaire or rating scale that would be completed by an expert either as part of or independent of the solution process.

Diary

If time is not constrained, an auxiliary technique that might be employed is to ask the expert to keep a diary during which information relative to all problems solved or expert-type actions taken would be recorded by the expert in the diary. If used at all, this should be combined with any or all of the preceding techniques. Of course, the nature of the expertise might be compatible with the keeping of a diary; if the expertise utilizes non-verbal (e.g., psychomotor) factors, the value of the diary will be sharply reduced. Diary responses would be content analyzed in order to make inferences about underlying mechanisms. A diary is largely under the control of the diarist, so to secure productive entries, it is necessary to supply the diarist with specific instructions as to what should be recorded.

The diary is not a primary methodology. Unless the expert is firmly committed to the task, diary quality may degrade over time as the expert becomes bored with the task of writing entries. Instructions are needed concerning the nature of the material to be supplied.

Conference

Another technique that was alluded to in connection with checklist development is to bring a small group of experts together in a room and present one or more problems to them. As a joint effort, experts attempt to solve the problems and discuss the various factors affecting the solutions. An expert can serve as moderator to stimulate the discussion, observe the process, and keep it on track.

The advantage of this technique is that it secures data from a group of experts at the same time and allows their responses to clash with each other, smoothing out (or highlighting, if one is interested in this aspect) individual differences in expertise. This technique assumes that all experts are working on the same problem. An alternative version is to have each expert work separately on the same problem and to provide feedback to them about other responses. This Delphi technique (Dalkey, 1969) does not lend itself as well as the group conference to determine the mechanisms that an expert

uses, because this method requires each expert to work individually and silently. It is more appropriate when one is trying to achieve a consensus rather than investigate mechanisms.

Historical Approach

If historical documentation is associated with the nature of the expertise, it may be possible to analyze these documents in an effort to infer the mechanisms of expertise. Examples are medical diagnoses, autopsy reports, test data, etc. This technique must be considered purely auxiliary, since the documentation may have only an indirect relationship to an expert's mechanisms. Nevertheless, this is an incidental data source not to be overlooked.

Tests

Expertise has occasionally been investigated in terms of the qualities that are supposed to underly that expertise. For example, one might hypothesize that figure/ground perception (see Koffka, 1935) is fundamental to a particular knowledge domain. One might then construct reversible face illustrations and test a group of experts to determine whether significant differences in the ability to perceive figure/ground reversal are associated with the expertise.

This is a very indirect way of securing expertise information and is not recommended except in a purely research situation. To utilize the technique one must have previously researched the knowledge domain, which may require more time than is available to the knowledge engineer.

In this connection it is necessary once again to distinguish between an expert's mechanisms and the information he or she can provide. The former are responsible for producing the latter, which is essentially the output of the mechanism. An example may illustrate the difference: Assume that the capability to make figure/ground reversals readily is the mechanism responsible for a certain type of expertise. An expert describes what he or she does as "examining a specimen to see what stands out in the foreground." This statement may or may not be sufficient for expert system development, but in any event it inadequately suggests the figure/ground reversal mechanism. The point is that expert system development may not have to dig so deep as an expert's mechanisms require. However this point must be studied empirically; it cannot be answered *a priori*.

Teaching

Another technique that can be used, either separately or as a part of a set of techniques, is to ask experts to try to teach their skills to novices. The novice can be the knowledge engineer, in which case the engineer introspects as part of the data collected, or it can be a third party, in which case the knowledge engineer observes the training process while recording data.

This technique is a variant on the interview method in which experts are asked to explain the principles underlying their actions. The interview is, of course, a tutorial situation as well as an explanatory process, and to that extent there is little difference between the two techniques. However, giving an expert the formal burden of instruction requires demonstration of the skills being taught, which tends to sharpen the information communicated. The expert could be asked to develop the test problems.

The difficulty with the tutorial approach is: How long does it go on? If it is genuine instruction, then it must continue until the student has achieved some specified skill level. The time required to achieve this level could be prolonged. Nevertheless, the effort to teach intelligibly often sharpens the teacher's perceptions of the knowledge domain.

Comparison

The most common method of researching expertise mechanisms is to compare the performance of two groups, one of experts, the other of nonexperts, in the solution of problems requiring some amount of expertise. The difference in behavior manifested by the two groups presumably reflects the operation of expert mechanisms. The same technique might perhaps be used in eliciting the rules utilized by experts. The idea is that the comparative technique would be used not in a research mode but rather applied to a knowledge domain for purposes of securing information. This technique has the disadvantage, however, that it requires considerable preparatory work before it can be implemented. Moreover, in comparative research on expertise, the emphasis is less on the derivation of consciously applied rules for procedures than it is on deep-seated mechanisms that raise, let us say, the average performer to one of great skill.

The application of mental models is a case in point; it may well be that the ability to apply a mental model or utilize photographic memory is what distinguishes a world master in chess, for example, from a chess player who is good but not a master. Whether one could make use of such almost unconscious mechanisms in the development of an expert system is an unanswered question. The essential quality of expertise may not be in the relatively superficial set of procedural rules that can be easily raised to consciousness; these are, indeed, essential, but they may not be sufficient to explain what makes experts perform as they do. If one takes seriously the statement that the expert system attempts to replicate an expert's mental processes, the effort may be foredoomed to failure if part of that expertise to be replicated consists of profoundly unconscious mechanisms. This is particularly the case when the expertise is less cognitive (e.g., in sports, music, art). The elicitation of information about an expert's expertise is an attempt to proceduralize expertise mechanisms and, indeed, if one cannot do this, the entire effort is valueless; it may not be possible to proceduralize some expertise mechanisms.

Experimentation

If one has some concept of the variables functioning in an expert's solution of problems, it might be possible to develop a set of problems that vary systematically in terms of these prespecified variables or dimensions. If these latter are actually relevant to an expert's expertise, his solutions or solution strategies should also vary in a systematic way with the variations in problems. To the extent that they do, this verifies that the variables/dimensions are important factors affecting the expertise. Of course, those variables/dimensions may not be easy to vary systematically, depending on their nature. The development of test problems varying along certain dimensional continua presents difficulties of some magnitude and would, of course, have to be left to someone with the requisite expertise. The very act of developing such tests varying along certain dimensions could itself be a means of eliciting expertise mechanisms from the test developer. This technique may, however, require more time and trouble than the knowledge elicitor has or wants.

Critique

A standard part of the process of eliciting information from an expert is to require the expert to critique the knowledge engineer's efforts to proceduralize the expertise mechanisms. It is common practice after translating the elicited information into a set of heuristic statements to take these back to the expert and ask how well they represent what the expert knows of his or her expertise mechanism. Experts are used throughout the expert system development process, sometimes as subjects for developmental tests of prototype software, often as critics of that software. Experts can also be used as subjects in acceptability tests (see Meister, 1987), or as critics of other experts' solutions to test problems and other material elicited from them. As was indicated previously, experts can be utilized as knowledge engineers in eliciting the information from other experts (e.g., each expert acting as subject interviewer as well as functioning as a subject).

A variant of this is to give experts solutions to various problems, solutions developed by other experts, and have them analyze these solutions, indicating where they agree and where they do not, and verbally giving reasons why.

Policy Capturing

An expert's judgmental policy can be captured to the extent that one can predict an expert's actions from the known characteristics of the stimuli to be evaluated. For example, if we ask physicians to rank 10 diseases in terms of overall virulence, four factors enter into that judgment: incidence of the disease, epidemiological distribution, duration of illness, and effects after recovery. If we can predict the overall ranking by a particular judge, using his or her judgments on these four factors, we would have captured that judge's policy in terms of the factors he or she considered most important.

It is not particularly germane to describe the details of the methodology (for which, see Christal, 1968). The important point is that the technique has been applied and has held up. It can be used to evaluate experts' mechanisms, but in order to do so it is necessary first to conceptualize the variables that go into that expertise. (This implies a sort of Catch 22, because if one can conceptualize those variables, is it necessary to go further? The answer is that one may be able to conceptualize all the variables that might comprise expertise in a particular domain, but not all or even most of them may be the ones that the expert actually uses.) This technique, like others cited previously, demands some study by investigators before it can be utilized.

A Final Note

One can be ingenious in finding variations, but all the techniques described depend on certain basic methods: interview, observation, questioning, critique of protocols, and pre-developed designs/solutions. No single variation is likely to produce significantly more than any other. The limiting factor is the inherent covertness of the mechanisms under study.

The problem of eliciting from an expert as much useful information as possible is a legitimate research topic. The goal of replicating an expert's mental processes (implicit in the effort to elicit an expert's information) for incorporation into the expert system is not a viable one, because it cannot presently be achieved. It is a legitimate research goal but not a developmental one.

There are several reasons why this goal cannot be achieved in system development. First, even if an expert's mental processes could be retrieved (and we have seen that unconscious mechanisms pose grave difficulties for us), the computer languages available to us (whether stochastic, deterministic, control-theoretic, or cognitive) are probably not effective to describe other than comparatively simple mechanisms. Second, it would be impossible to determine when expert mechanisms had been correctly captured in precise detail, because there is no external criterion for this determination. One might ask a group of experts to judge this point, but who could guarantee that they could do so, even if there were a consensus, because all experts might be similarly limited in their capacity to judge.

From a system development standpoint, one should ask whether the replication of mental processes is a necessary goal in expert system development. Certainly one wishes to gather as much relevant information as possible, but the attempt to replicate mental processes could lead to excessive and perhaps unnecessary efforts to achieve a perfection not actually required or feasible. One fact is comforting, however. Although one can never know whether an expert's mental processes have been adequately elicited and replicated, one can determine on the basis of testing whether the expert system is reasonably effective.

Finally, purely analytic efforts such as the one leading to this report will never solve the problem. Only empirical research on the dimensions of expert mechanisms will do so.

CONCLUSIONS

1. Expertise mechanisms must be differentiated from procedural rules of thumb; the former are extremely difficult to elicit.

2. There is no external criterion to specify when one has elicited a sufficient or optimal amount of information from the expert; all one can do is build the expert system and test its effectiveness.

3. Ways of eliciting information from experts are constrained by the fact that the knowledge they possess is covert, in many cases lacks an external criterion, and relies on the experts' own awareness of the expertise mechanisms.

4. It is recommended that in eliciting information from experts, the knowledge engineer make as much use as possible of other experts in the knowledge domain and that a variety of techniques (e.g., interview, problem-solving, verbalization) be applied concurrently. Empirical studies of expertise elicitation should also be conducted.

REFERENCES

Adelson, B. (1981). Problem solving and the development of abstract categories in programming languages. Memory and Cognition, 9, 422-433.

Adelson, B. (1984). When novices surpass experts, the difficulty of a task may increase with expertise. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 483-495.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Anderson, J. R., Greeno, J. G., Kline, P. J., & Neves, D. M. (1981). Acquisition of problem solving skill. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Antonelli, D. (1983). The application of artificial intelligence to a maintenance and diagnostic information system (MDIS). Proceedings, Joint Services Workshop on Artificial Intelligence in Maintenance, Los Angeles, CA.

Bhaskar, R., & Simon, H. A. (1977). Problem solving in semantically rich domains. Cognitive Science, 1, 193-215.

Biggs, S. F. (1978a). Cognitive structures in the comprehension and memory of computer programs. An investigation of computer program debugging (Report AD-A060 522/0). Washington, DC: National Technical Information Service.

Biggs, S. F. (1978b). An empirical investigation of the information processes underlying four models of choice behavior. In T. J. Burns (Ed.), Behavioral experiments in accounting, II. Columbus, OH: Ohio State University College of Administrative Science.

Boose, J. J. (1986). Expertise transfer for expert system design. Amsterdam, The Netherlands: Elsevier.

Bouwman, M. J. (1978). Financial diagnosis: A cognitive model of the processes involved. Unpublished doctoral dissertation, Carnegie-Mellon University, Pittsburgh, PA.

Bouwman, M. J. (1983). Human diagnostic reasoning by computer: An illustration from financial analysis. Management Science, 29, 653-672.

Brown, A. L. (1978). Knowing when, where and how to remember: A problem of metacognition. In R. Glaser (Ed.), Advances in instructional psychology (Vol. I). Hillsdale, NJ: Erlbaum.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. Cognitive Psychology, 4, 55-81.

Chi, M., Glaser, R., & Reese, F. (1981). Expertise in problem solving. In Advances in the Psychology of Human Intelligence (Vol. I). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Feltovich, R. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Christal, R. E. (1968). Selecting a harem--and other applications of the policy capturing model. Journal of Experimental Education, 36, 35-41.

Dalkey, N. C. T. (1969). The Delphi method: An experimental study of group opinion. Santa Monica, CA: Rand Corporation.

Davis, R., & Lenat, D. B. (1982). Knowledge based systems in artificial intelligence. New York: McGraw-Hill.

de Groot, A. D. (1965). Thought and choice in chess. Paris: Mouton.

Elstein, A. A., Shulman, L. W., & Sprafka, S. A. (1978). Medical problem solving: An analysis of clinical reasoning. Cambridge, MA: Harvard University Press.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner and A. L. Stevens (Eds.), Mental models. Hillsdale, NJ: Erlbaum.

Glaser, R. (1984). Education and thinking. American Psychologist, 39, 93-104.

Glaser, R. (1985). Thoughts on expertise (Technical Report No. 8). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Jeffries, R., Turner, A. A., Polson, R. G., & Atwood, M. E. (1981). The processes involved in designing software. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Johnson, P., Hasebrock, F., Duran, A., & Moller, J. (1982). Multimethod study of clinical judgment. Organizational Behavior and Human Performance, 30, 201-230.

Kahn, G. Nowlan, S., & McDermott, J. (1985). MORE: an intelligent knowledge acquisition tool. Proceedings, Ninth International Joint Conference on Artificial Intelligence, Los Angeles, CA.

Kelly, G. A. (1955). The psychology of personal constructs. New York: Norton.

Koffka, K. (1935). Principles of gestalt psychology. New York: Harcourt Brace.

Larkin, J. H. (1981). Enriching formal knowledge: A model for learning to solve textbook physics problems. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Larkin, J. H., McDermott, D., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.

Lesgold, A. M., Feltovich, P. J., Glaser, R., & Wang, Y. (1981). The acquisition of perceptual diagnostic skill in radiology (Technical Report PDS-1). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Lewis, C. (1981). Skill in algebra. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Malhotra, A., Thomas, J. C., Carroll, J. M., & Miller, L. A. (1980). Cognitive processes in design. International Journal of Man-Machine Studies, 12, 119-140.

Marcus, S., McDermott, J., & Wang, T. (1985). Knowledge acquisition for constructive systems. Proceedings, Ninth International Joint Conference on Artificial Intelligence, Los Angeles, CA.

McKeithen, K., Reitman, J. S., Rueter, H., & Hirtle, S. C. (1981). Knowledge organization and skill differences in computer programmers. Cognitive Psychology, 13, 307-325.

Meister, D. (1987). Behavioral testing and evaluation of expert systems. Proceedings, Second International Conference on Human-Computer Interaction, Honolulu, Hawaii.

Newell, A., & Simon, H. A. (1972). Human problem solving. New York: Prentice-Hall.

Reitman, J. S. (1976). Skilled perception in Go: Deducing memory structures from inter-response times. Cognitive Psychology, 8, 336-356.

Schneiderman, B. (1977). Measuring computer program quality and comprehension (Technical Report No. 16). College Park, MD: University of Maryland, Department of Information Systems Management.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cooke, N. M., Tucker, R. G., & De Maio, J. C. (1985). Measuring the structure of expertise. International Journal of Man-Machine Studies, 23, 699-728.

Scribner, S. (1984). Studying working intelligence. In B. Pogoff and J. Lave (Eds.), Everyday cognition: Its development in social context. Cambridge, MA: Harvard University Press.

Shaw, M. L. G. (1982). PLANET: Some experience in creating an integrated system for repertory grid applications on a microcomputer. International Journal of Man-Machine Studies, 17, 345-360.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, NJ: Erlbaum.

Sternberg, R. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Erlbaum.

Vitalari, N. P., and Dickson, G. W. (1983). Problem solving for effective systems analysis: An experimental exploration. Communications of the ACM, 26, 948-956.

Voss, J. F., Green, T. P., Post, T. A., & Penner, B. C. (1983). Problem solving skill in the social sciences. In G. Bower (Ed.), The psychology of learning and motivation: Advances in research theory. New York: Academic Press.